Appendix 2

Application of Genetic Markers to Forest tree species

Draft report to IPGRI of the project "Developing Decision-making Strategies on Priorities for Conservation and Use of Forest Genetic Resources"

by Dr. Gene Namkoong and Dr. Mathew P. Koshy (University of British Columbia)
2001

1. Introduction

The application of genetic markers such as allozymes and DNA (Deoxyribose Nucleic Acid) markers has shown to be useful in studying genetic diversity in humans, animals and plants (Cruzan, 1998). This section of this manual attempts to give guidelines for choosing appropriate molecular tools for the study of genetic variation in species under different conservation status. In general, forest trees can be categorized as totally wild species , trees of economic importance (both un- or semi- domesticated species) and lastly, model forest trees, which are being intensively managed through tree breeding.

2. Application of genetic markers to wild species

Usually, very little information is available to design a genetic study for a wild species. Therefore, a preliminary study should be carried out in order to guide the choice of the most suitable technique and enable an estimate of the associated costs per sample. Unknown pitfalls can also be detected while carrying out the pilot study.

2.1. Pilot genetic study

To begin genetic research on a tree species, a set of basic information ought to be known, in particular, its ecology and demography. Their mode of reproduction, and mating systems are also important aspects to be understood. In the pilot genetic study, where a subset of individuals are sampled from a number of populations across the range of the species, a small sample size of 5 to 8 individuals per population would be enough to test genetic diversity within a population. This will give sufficient information for hypothesizing population structure and compare genetic diversity between populations. Alternatively, there could be fewer populations ($k = 4$) with many more individuals ($n = >20$) per population.

2.2. Collection of samples

The types of tissues required for genetic marker study would be based on the genetic question being asked and the technique of choice. If little to no information about genetic diversity within or between populations is known, isozymes can be an effective genetic marker. Many plant species have been well studied and characterized using isozymes (Tanksley and Orton, 1983; Mitton, 1983). For these studies, fresh actively growing leaf samples will be the easiest material to use. However, other plant parts also can be used. Sprouting seedlings, and active meristems will also be ideal as enzymatic activity in these tissues will be higher compared to dormant tissues. If newly swollen seeds or germinating seedlings are used, isozyme bands are sharper and cleaner which greatly enhance scoring of the gel. For DNA marker studies, the quantity of tissue could be critical for the choice of marker, particularly when not using the technique Polymerase Chain Reaction (PCR) (Mullis, 1986, Saiki 1985). Needles, pollen, seeds, buds and fully mature seeds and possibly bark (in particular the xylem layer) could be used for DNA studies. Age of tissue is particularly important for the quality of DNA (De la Cruz, 1995, Jobes *et al.*, 1995).

The amount of tissues that are necessary for any particular type of genetic studies will depend mostly on the technique of choice. The quantity of tissue required for isoenzyme

analysis depends on the age of the tissue, the older the tissue greater the amount. Another source of tissue, which could be used for isozyme studies, is pollen (Protopapadakis, 1987). For DNA studies that are PCR based such as Randomly Amplified Polymorphic DNA (RAPD), small amounts (1-2 μg) of DNA would be enough. However, for techniques such as Amplified Fragment Length Polymorphism (AFLP) 5-15 μg of DNA per sample is required for restriction digestion of DNA prior to PCR reactions. DNA isolation protocol using CTAB solution yields on average from 20 to 100 ng/gram of wet weight while others (De la Cruz, 1995) have reported yield as high as 1.5 mg/g to 24 mg/g pending on species. Dried herbarium specimen could also be used to isolate DNA, particularly if they have not been preserved with any of the formalin family of chemicals (Rogers and Bendich, 1985).

Plant tissues should be collected during dry days if possible. If moisture is a problem, allow the plant material to air dry or dry with paper towel prior to storage. Moisture will enhance fungal growth, hence leading to degradation of plant material either for DNA or isozyme work.

Degradation of tissue for either DNA or for isozyme studies is a great detriment. Measures such as immediate cold storage (minimally 4°C to ideally -70°C) of the tissues while collecting in the field will greatly enhance the quality of DNA, while it is absolutely necessary for isozyme work. Even for seed collections, the seeds should be cleaned off all fleshy parts, dried and packaged air tight for cold storage (-20°C). Field collection tanks (eg. MVE ™, Cryogenics) with liquid nitrogen are ideal, to flash freeze and to keep the samples frozen during the time out in the field for up to 3 weeks. Dry ice in styrofoam containers can also be used when collecting samples, though, good sources of dry ice maybe difficult to locate in the field. If there is no cold storage facility available, plant materials should be quick dried in a mid. temperature oven providing the thickness of the material is no greater than 0.5cm. A very absorbent plant press may be another method of preservation of tree material for laboratory studies. Frequent changes, particularly in the early stages of drying, of absorbent material should be done to prevent rotting of plant material. A novel technique of using silica gel for preserving leaves from trees without the use of cold storage or a plant press could be considered when looking for an alternative method of collection (Chase et Hills, 1991).

Labelling samples correctly is crucial to any studies. Care should be taken to use waterproof markers and invisible tape to protect any writing on sample tubes. Microtubes (1.5ml) with secure caps make good vessels for collection of small amounts of tissues or seeds. Larger tubes (50ml) are useful for larger volumes of tissues.

2.3. Long term storage of samples

For long term storage, particularly for isozymes, plant tissue should be kept cold (minimally at 4°C to ideally -80°C) and dry. It is essential that the plant material be processed as quickly as possible. Cold storage of up to 3 months may yield high quality of isozyme but the activity level will steadily diminish. Beyond six months very little isozyme activity will be detected. For DNA studies, it is essential that the plant material be cry-o-preserved at -80°C or lyophilized and stored at room temperature. Since little degradation of DNA could occur at ultra-low temperature or from dried material, plant material can be stored indefinitely for DNA with the above two methods.

Seeds are ideal for long-term storage for any genetic studies. Once dried and kept cool (-10°C to -15°C) at low moisture level (6-7%), viability of the seeds remains very high (Pita *et al.*, 1998).

2.4. Techniques of choice

2.4.1. Isozymes

This is one of the most efficient genetic markers for assessing diversity (Lewontin and Hubby, 1966). In forestry, isozymes have been used to study genetic variation within and between populations, population structure, phylogeny, and to elucidate mating patterns among natural populations as well as experimental populations (Mitton, 1983, Hamrick and Godt, 1989, El-Kassaby and Ritland, 1996a). Pollen contamination in the in seed orchards can also be detected using allozyme markers (El-Kassaby and Ritland, 1996b). Isozyme data is also generated for the comparison of genetic variation within and between populations and towards deduction of mating systems in many tropical forest tree species (Loveless, 1992). The advantages of using isozyme markers are many folds. Many samples (20-30) can be processed at once and multiple isozymes (1 to 6) can be scored from the same individuals, using starch gel electrophoresis. Polyacrylamide, cellulose acetate and agarose gel over lay are also used as substrate to separate proteins. The cost per samples can be minimal pending on the enzyme and the substrate used to screen proteins ($0.50 to $2.00 per sample per loci). Individuals with homozygous genotype are usually indicated by one band, while heterozygotes may show two to five bands pending on the isozyme being scored. A screening of 20 to 30 individuals is useful for determining the number of common alleles for a given enzyme. The time required to learn the technique is relatively short and one can quickly screen many isozyme to assess genetic diversity (Heterozygosity, H). Other genetic parameters such as P (Percentage of polymorphic loci), F_{st} or G_{st} can also be calculated (Berg and Hamrick, 1997). One of the disadvantages of using isozyme is the quick degradation of proteins and enzymes. The chemicals used can be particularly carcinogenic (eg. MTT and PMS). The long hours, from the time of grinding up the samples to staining is another drawback (8-10 hours) when using starch gel electrophoresis. However, despite time saving when producing protein bands on acetate gel there could be artifacts due to their short resolution distance. Due to the light sensitive nature of the enzyme stains and the spreading of the proteins (particularly for starch gel system), the gels should be scored soon after staining.

For most isozyme researchers, scoring of the gels can be come an “art”. Keeping meticulous record, using photographs or preserving the gels for future comparison will help when scoring gels. When possible, different individuals from different populations should be assayed on the same gel aiding in the detection of different allelic forms. Strong knowledge of the substructure formation of the isozyme is critical when scoring complex (eg. tetrameric protein) patterns. A common convention for data collection is the use of digit such as 1, 2 or 3 to denote different alleles so for example, the notation of 1,1 and 2,2 would represent two homozgote alleles from two individuals whereas 1,2 would be the notation for a heterozygote individual where the individual would have both allelic form. This pattern will be scored for a monomeric isozyme (Murphy *et al.*, 1996). The calculation of genetic diversity within a population is based on the % of polymorphic loci where %P = <0.01 would be considered as a rare allele in large samples of over 100 and %P = <0.05 would be considered as a rare allele in small sample size of less than 100

individuals. Among population we can calculate heterogeneity tests based on allele frequencies at each locus using programs like Biosys-1 (Swofford *et al*. 1981) and the G-test (Sokal and Rohlf, 1981). A common measure of genetic distance to test for genetic differentiation is the use of Nei genetic identity. For a comprehensive review of distance measures reader should consult Nei (1987) and Swofford *et al*. (1996). For other genetic diversity programs that could be used see appendix of Swofford *et al* (1996) pp.510-514.

2.4.2. Random Amplified Polymorphic DNA (RAPD)

RAPD (Williams *et al*., 1990) is another genetic marker useful for estimation of heterozygosity in species of trees that are not widely studied. This technique uses PCR with very small pieces of (10 bases or shorter) primers (oligonucleotides). Since these oligos will hybridize randomly on any given genome, it produces many loci for any one individual. It is possible to have as many as 3 to 8 loci per agarose gel. Ethidium Bromide (EtBr.) is used to stain DNA fragments. The advantage of this technique is the simplicity of the system and the relatively low cost ($2.00 per sample per oligo used). Normally, one can screen up to 40 samples per agarose gel with one oligo and with thousands of combination of oligos, one can quickly screen as many loci as required for the calculation of diversity or for phylogeny purposes.

The disadvantages of using RAPD technique are the faint bands due to the nature of PCR and the repeatability of the PCR products. To avoid inconsistency of data, it is important to start with very clean template DNA, using quantified concentration, and consistent methodology (Ayad *et al*., 1995, Karp *et al*., 1997). The DNA bands from one individual do not necessarily correspond to the next individual unless probing of the bands have been verified using Southern Blot Technique (Southern, 1975). Running the same samples with the same oligomer on different days would help to determine the repeatability of the markers being scored. Same electrophoresis conditions and concentration of agarose gel are essential to maintain consistent results when scoring RAPD bands. The use of Ethium Bromide (EtBr) requires great care since EtBr is a known carcinogen.

Since RAPD data is scored as a dominant marker, there are several concerns about the usage of the data mostly due to problems unique to RAPD markers. Consistency between gels of the same size bands appearing at the same intensity (little shadow or pale bands) cannot be assured across the entire project. Slight changes of substrate to the PCR reaction will alter the bands produced by PCR. It is ideal if RAPD data is collected for a pedigree whenever possible or the usage of known fingerprint is used in every gel given this individual is highly repeatable and robust throughout the whole experiment. Data is scored as 0 or 1 (absence or presence of bands across individuals) for a given locus. Ideally, size markers are chosen to give landmarks for any given loci. Genetic diversity based on shared fragments between samples has been calculated for RAPD data. Clark and Lanigan (1993) and Kremer *et al* (1998) have used Lynch and Crease's model (1990) to calculate differentiation within and among populations. For other genetic diversity programs that could be used see appendix of Swofford *et al* (1996) pp.510-514.

2.5. Additional considerations on the different techniques available

ISOZYMES. The use of protein data as useful genetic marker is recommend to this day (Sunnucks, 2000). The direct comparison of different studies using isozyme is very still

common and the reliability of the data is not often questioned. With this technique, the allele genealogy is rarely feasible and a limited number of loci is taken into consideration, therefore PCR markers such as microsatellites have taken the place of isozyme markers even at a high cost per data point (for more details, see paragraph 3 on trees of economic importance). However, the base line for estimation of genetic diversity is still built on isozyme data that has been produced in the past, and on more rarified current investigations carried out on species with poor genetic characterization.

RAPD. "Band sharing" or co-migration of a given band is used for data analysis. The band sharing index, based on the principle of band similarity, i.e. bands which have certain molecular weight range, is used when calculating relatedness of a given set of individuals.
Test for independent loci are required to avoid allelic pairs and tight linkage groups. Naturally, allele genealogy is not possible since short primers can complement in different areas of the genome and yet produce similar size PCR products. The use of RAPD data when comparing different studies (mainly from different laboratories) are rather limited since the origin of these bands are tenuous and chosen primers will need to be tested on the same individuals, population or species to make any type of direct comparison.

2.6. Summary on the applications of the techniques described

Isozyme genetic markers have been used for many decades and are still one of the most reliable and efficient tools in the estimation of heterozygosity and predication or confirmation of mating system for mammalian or plant species. This technique should be chosen over most DNA techniques, since it is so well studied and cost efficient. It is useful for pilot studies as it reveals whether a population is highly variable and also detects variability across different population. It is of great use in the elucidation of genetic information for relatively unknown species and it has been used to determine species delimitation and possible hybrid zones (Rajora, O, 1989). In tropical forests with high species diversity, information on species population biology and genetic structure are critical for the conservation of species and their genetic variability. The same information is also fundamental for the selection of species potentially suitable for domestication and cultivation.

3. Application of genetic markers to tree species of commercial importance

3.1. Collection of samples

With respect to trees of economic importance, germplasm banks may be considered as important genetic resources for the future. Both tissue and seed banks are required to store valuable genetic material that could be used when the need for new hybrids strains arise and the need for a new gene pool are required for reforestation. This becomes even more important when habitats have been destroyed and the need to restore the original forest species is required. Great care should be taken when choosing particular populations that contain in their genetic makeup adaptive mechanisms to different environmental conditions. With the advent of genetic transformation technology, the

potential for new clones of tree species is a reality. With current regeneration technology and molecular transformation capabilities, germplasm banks will be a rich source of genetic material. Herbarium plant collections that have not been treated with formaldehyde are potential material for DNA work. Ideally, viable seed collections will ultimately be the best source of plant material for current as well as future research. Ongoing germination experiment should be done to test the viability of seeds particularly when long-term storage (>5 years) is desired to monitor genetic diversity. Aside from the previous two techniques already mentioned, we will examine the following genetic techniques for the semi-domesticated and domesticated trees, which are mostly based on North American trees and the applicability to tropical tree species.

3.2. Techniques of choice

3.2.1. Isozymes

Similarly to the case of wild species, for those species of commercial interest isozymes would represent the first technique of choice to monitor population structure, colonization, and mating system (Cruzan, 1998). In North America, studies of colonization of *Pinus contorta* (Aitken and Libby, 1994), range survey of Douglas Fir (El-Kassaby and Ritland, 1996a, Li and Adams, 1989, Yeh and O' Malley, 1980), ecological influences on genetic variability of Douglas Fir (Moran and Adams, 1989), and mating system of natural and shelter-wood stands of Douglas fir (Neale and Adams, 1985) have all been studied using allozymes. Speciation and systematics of various trees species such as *Pinus* (Millar *et al*, 1988), *Populus* (Rajora, 1990), *Larix* (Liu and Knowles, 1991) have also been studied using allozymes. Hamrick and Murawski (1991) have investigated uncommon neotropical tree species using isozymes. A gradual shift towards DNA markers would most likely reduce the use of isozymes. Nevertheless, when funding and time are limited isozymes are still preferred.

3.2.2. RAPD (Random Amplified Polymorphic DNA)

DNA markers such as RAPD can be very useful when studying genetic diversity. Using a minimum amount of equipment, a PCR machine, agarose gel eletrophoresis system and UV viewer for the agarose gel, one can obtain many types of loci that are used in the determination of genetic diversity, taxonomic identifications, and paternity analysis (Van de Ven and McNicol, 1995, Kamalay and Carey, 1996, Nesbitt *et al*. 1995. When two of these 10 nucleotide base primers are positioned, when duplication of the genomic template might occur, unique fragments will be generated, which can be separated on agarose or low concentration polyacrylamide gels.

Once a particular band (functioning as a locus) useful in differentiating between clones or even species, is identified, it can serve as a constant diagnostic marker for future screening. These markers are known as SCAR (Sequence Characterized Amplified Region) (Bodenes *et al*. 1997). Ideally loci are cloned and sequenced to determine the first 50-100 bases of both ends of the fragment. Then primers are made either with the original RAPD sequence or sometimes with new sequences. These SCAR markers are very useful to observe specific allelic differences. Other methods such as Southern (1976) Blots can be used to verify the variable locus observed with the RAPD maker, as verification of the same locus will give more confidence to RAPD data.

Random Amplified Microsatellite Polymorphism (RAMPO), which is a RAPD based technique, is a more reliable method than RAPD and it gives simpler bands. This technique involves amplifying loci with a "classic" 10 base oligonucleotide and a microsatellite anchored oligonucleotide (MP-PCR). Subsequently, the agarose or modified polyacrylamide gels will be blotted and probed with oligonucleotide probes $(CA)_8$ or $(GA)_8$. The loci that are observed can be species specific and different alleles can be spotted at a locus. RAMPO was used in the study of the taxonomy and phylogeny of the Guinea yam (Ramser *et al*. 1997).
Recently, more studies are using RAPD technique for generating genomic maps (Tulsieram *et al*, 1992). Specific makers for certain traits such as blister rust resistance (Devey, *et al*, 1995) and inheritance of terpene compound have been mapped using RAPD markers (Plomion *et al*, 1996).

3.2.3. Microsatellites

Microsatellites, defined as tandemly repeated sequences whose unit of repetition is between 1 to 9 base pairs, have quickly replaced isozyme as the choice genetic marker for population studies. If the laboratory can afford the added cost of equipment and the initial investment to obtain the microsatellite primers (see Karp *et al*, 1997 for isolation and detail explanation of the technique), these markers are preferred to isozymes. Microsatellite will be able to provide many more alleles per locus than isozymes. The rate of mutation can be as high as 10^{-4} / gamete / generation (Edwards *et al*. 1992). Due to the pattern of inheritance, it is much easier to score a few loci (5-20) with many alleles than many isozyme loci (10-50) with a few alleles. These markers are codominent, PCR based, and are much more reliable than RAPD markers. Since nucleotide repeats occur in all organisms, these can be ideal Mendelian genetic markers for studying population structure, migration patterns, and mapping of genomes. Once species-specific microsatellite primers are found, there could be signature allelic patterns for separating out different populations and even identify individuals within a population. Naturally, heterozygosity levels, immigration, and rates of mutations can be studied using this type of genetic markers. In humans, there are known heritable diseases that are direct results of microsatellite changes, for example, fragile X and Huntington's syndrome (Sutherlands *et al*, 1995). Interestingly, both syndromes are due to triplet repeats. By far the most common of the microsatellite repeats are dinucleotide repeats. For animals, these are usually CA repeats whereas for plants they are either TA or GA repeats. This is not to say that microsatellites are not neutral markers, in fact, many of these markers appear to be selectively neutral founded predominately in introns (Jarne and Lagoda, 1996).
One of the major drawbacks of this technique is the high cost involved in the development of the repeats (CDN $5,000-10,000 for cloning and screening of clones, CDN $5,000-10,000 for sequencing and screening primers for potential polymorphic sets of microsatellites primers). Another drawback is associated with the fact that microsatellite primers are usually unique at the species level and cross-species amplification does not usually lead to desirable results. This could be due to the biased skewing of polymorphic loci for the candidate species rather than for the related species (Amos and Rubinsztein, 1996). Unless, the species are closely related (i.e. separated by no more than a few thousand generations), microsatellites are not useful markers for phylogenetic studies beyond the taxonomical level of orders or families. Since conifers

have such large and complex genomes, the isolation of these primers can be labour intensive and costly. Once the primers are developed, collection of the data for any given population is very straight forward, particularly, if an automated sequencer is available. The scoring of these alleles can be tricky at times especially if there are only a few (<6) loci to chose from. It is recommended that many more loci be isolated so that a few loci could be discarded, hence gaining reliability. Many of complexities in scoring of microsatellite data are due to the nature of the template as well as how the microsatellite primers were designed. This is particularly difficult for cross species amplification.

Many cultivated species such as rice, soybean, sugar beet, wheat, yam and barley have been well studied using many genetic markers such as allozymes and RAPD but with the advent of microsatellites, these plants also have microsatellite markers added to their collection of genetic data (Wu and Tanksley, 1993; Powell *et al* 1996; Schmidt and Heslop-Harrison, l996; Ma *et al*, 1996; Terauchi and Konuma, 1994; Becker and Heun, 1995). Currently, many forest geneticists are using and searching for microsatellites on conifer species. In species such as *Picea abies* and four species of *Pinus* (*radiata*, *strobus sylvestris* and *taeda*) CA and GA repeats are abundant (Pfeiffer *et al*, 1997; Smith and Devey, 1994; Echt and May-Marquardt, 1997; Echt *et al*, 1996; Kostia *et al*. 1995).

Initial screening for microsatellites in tropical trees has shown abundant CA and GA repeats amongst five trees studied (Condit and Hubbell, 1991). Additionally, in *Dryobalanops lanceolata*, which is a large canopy tropical tree, bark as well as leaves was assessed for any genotypic difference within an individual. The same genotype was found for both the leaves and cambium tissues on collections from different region of the tree (Terauchi 1994). This information is particularly useful for tropical trees as the collection of leaves from tall canopies could be particularly difficult. DNA isolated from bark can also be used for investigation in conifers. Chase *et al* (1996a, b) estimated gene-flow distance and identified of paternal genotypes for 72 mating events using 5 loci. Chloroplastic microsatellites have been identified using total chloroplast sequences from rice, tobacco and liverwort (Vendramin *et al*, 1996). Powell *et al* (1995) studied potential mono repeats of microsatellites in chloroplast and found mainly A and T repeats. Inheritance of chloroplast is through maternal lines in angiosperm whereas in conifers they are mostly inherited through paternal lines. Specific chloroplast microsatellite primers (20) were designed by using total black pine chloroplast sequence (Vendramin *et al*, 1996). These primers seem to be fairly universal. In other conifer species (*Abies, Cedrus, and Picea*) it gives reduced success (100% to 75%) as the genera become less and less related to *Pinus*. The application of these primers to natural populations of *Abies* have revealed larger than expected number of detected haplotypes which could indicate high levels mutations rate for some of the chloroplast loci (Vendramin and Ziegenhagen, 1997).

3.2.4. AFLP (Amplified Fragment Length Polymorphism)

Amplified Fragment Length Polymorhpism (AFLP) has been extensively used in the plants (Vos *et al* 1995). Aside from two known species of trees (Eucalyptus and coconut) and a critically endangered angiosperm (*Astragalus cremnophylax*), all other genetic studies using AFLP markers have been carried out on cultivars like wheat, barley, melon, rice, corn and soybean. The study of developmental genes in *Arabadiposis* has taken advantage of this new technique (Cnops *et al*, 1996). For many of the cultivars, AFLP is

used to create extensive linkage maps that have already been generated from RAPD and RFLP (Restriction Fragment Length Polymorphism) markers. Since none of these markers produce large number of loci for any one reaction, AFLP markers are preferred in linkage mapping. AFLP markers are generated using two restriction enzymes (a 6-base pair cutter and a four-base pair cutter often referred to as rare and frequent cutters respectively). Then specific restriction site adapters are linked to the sticky ends of the fragments of DNA. When using specific restriction enzymes and subsequently specific adapter primers which have random bases chosen at the 3'end of the primer, only certain fragments will be amplified using PCR (for further details see Karp, 1997, Ritland and Ritland, 2000). More and more specific nucleotides can be used to restrict the DNA pool for fragment amplification. These are designated as +1, +2, +3 and even +4 nucleotide, chosen at either end of the primers. This technique does generate dominant markers similar to RAPD, which are scored as present or absent alleles. Yet these markers are highly repeatable and using as few as 4 to 10 primer combination, many loci (40 to 200) can be scored. They are easier to use than the co-dominant markers generated by RFLP (Restriction Fragment Length Polymorphism). Another advantage is the high rate of polymorphism over RFLP and many more loci can be scored than a RAPD primer (10-20 [AFLP] verses 4 [RAPD] per primer set).

The disadvantage of AFLP compared to RAPD is the need for larger amount of DNA (10 µg verses 10 ng per sample). The quality of DNA is of concern, since restriction enzyme could cut irregularly or produce unreliable results. The fact that AFLP markers are dominant is a disadvantage over RFLP markers. Despite these drawbacks, AFLP markers are becoming popular DNA markers for studies on genetic diversity (Travis *et al*. 1996, Russell *et al*, 1997, Zhu, *et al.* 1998, Tohme *et al.* 1996), estimation of outcrossing rates (Gaiotto *et al*, 1997, Marsan *et al*., 1998), linkage mapping/alignment of maps (Keim *et al*, 1997, Waugh, *et al*, 1997, Becker *et al*. 1995, Qi and Lindhout, 1997, Mackill, *et al.* 1996, Rouppe van der Voort *et al.*, 1997), paternity analysis (VanToai *et al*, 1997), dense mapping of a particular area on the genome (Meksem *et al*, 1995) and mapping to a chromosome for location of a quantitative or inheritable trait (Pakniyat *et al.*, 1997, Nandi, *et al*. 1997, Cnops *et al.*, 1996). Other studies such as the identification of genotype for tracking the origin of domestication in wheat (Heun, *et al*. 1997) have shown that AFLP markers are plentiful, highly repeatable and reliable.

These markers have been used in tree species to estimate outcrossing rate. Results obtained with these markers were compared with those derived from isozymes. Gaiotto *et al.* (1997) used both RAPD and AFLP markers to re-estimate outcrossing rates in *Eucalyptus urophylla* and have found that both markers generated rates that were similar in magnitude to those of isozymes. AFLP was used to study variation between tall, intermediate, and dwarf forms of *Cocos nucifera* (Perera *et al*. 1998). AFLP is currently being used in other species such as Lobolly Pine for the creation of a genomic saturated marker map (Remington *et al*. 1999). The usefulness of this technique for new taxa of tropical trees is worth considering (cost per sample/ per data point/ $3.50 to $5.00 pending on restriction enzyme use as well as the method for labeling the fragments to visualize the bands).

3.2.5. Organelle DNA markers

Chloroplast organelles are predominantly inherited through the paternal line in conifers, and though the maternal line in the angiosperms. Pollen flow can be effectively studied using chloroplast markers. By eliminating the influence of the paternal parent from the chloroplast, the effect of the maternal protoplasm can be carefully monitored. All members of the Pinaceae family have paternally inherited chloroplast genome, where as mitochondrial genome are inherited through the maternal line (Ali *et al*. 1991; Neale *et al*. 1989, 1991; Sutton *et al*., 1991a,b Szmidt and Hallgren, 1987). Taxodiaceae and Cupressaceae, have paternal chloroplastic and mitochondrial inheritance (Neale and Sederoff, 1989 and Neale *et al*., 1986).

Introgression zones for *Picea* species have been studied using RFLP data on chloroplasts. El-Kassaby and Szmidt (1988) have used this methodology for tagging and identification of pure and mix seedlots of *Picea sitchensis, P. glauca* and *P. englemanni*. Subsequently, the most variable bands were cloned and used as probe for quick screening of various seed lots (Sutton *et al* 1991a; Szmidt *et al* 1988). Unlike the study of point mutation (using sequence data) between individuals, the use of RFLP/Cp. markers is limited to the genus level of diversity (Ali *et al*., 1991). PCR specific chloroplast primers, particularly for the highly variable regions, would eliminate the need for Southern (1976) blots and probing.

Universal primers for three intron regions on the chloroplast that are flanked by trnT and trnF have been designed for use on most plants including gymnosperms (Taberlet *et al.*, 1991). Using these makers, the study of post glacial migration routes of oak have shown a 13 base pair repeat that differentiate populations of East Anglican oaks from those found in other parts of Britain or Europe (Ferris *et al*., 1995). In total there are 16 chloroplast primer sets and 12 mitochondrial sets (Demesure *et al.*, 1995; Dumolin-Lapegue *et al*., 1997). These primers are used to amplify various plant genomes and subsequently using restriction enzyme these fragments are digested to obtain RFLP data. This method is quick and less labour intensive than the traditional RFLP methods (8 hours verses 3 days). Effectively the cost is somewhat lesser ($1.00 verses $1.80 per sample). The latest PCR based chloroplast markers are the universal microsatellite primers (Vendramin *et al.*, 1996).

3.2.6. Other molecular genetics markers

Together with PCR based fragment analysis, DNA probes such as RFLP have been used to create genome maps. Other DNA probe techniques such as VNTR (Variable Number Tandem Repeats) have been used extensively in studying diversity particularly in humans and other mammals. We have already described microsatellites, a subgroup of VNTR.

The other subgroup is known as minisatellites. They have repeats that vary between 16 and 64 base pairs. Minisatellites are long, therefore southern blots of genomic DNA, restriction digested, are probed with these minisatellites (Jeffreys *et al.*, 1985). Certain types of probes (33.6 and 33.15) (Jeffreys *et al.*, 1985) have been used to study polymorphic levels in various plant genomes (Sharon, *et al.*, 1995, Rogstad *et al.*, 1988). Another source of probes is based on a bacteriophage, M13. This probe has been known to detect polymorphism in *Picea glauca, Pinus torreyana* and *Populus* species (Rogstad *et al* 1988). Although polymorphism has been detected with these VNTR methods, the labour intensiveness of this method and the necessity to use radioactive labels limit their

uses. Up to now, no (at least to the knowledge of the author) studies have extended the use of VNTR for scoring large populations of forest trees to calculate genetic diversity or for bio-systematic purposes.

The co-dominant inheritance of nuclear RFLP markers has made them useful for the construction of genetic maps in conifers (Neale *et al* 1992). The concept of RFLP mapping for genome was first described by Botstein *et al*. (1980). In plants such as rice, lettuce and maize genome maps have been made with RFLP markers. They served as the base map for other molecular markers including isozymes, RAPD and SSR and AFLP (Nandi *et al*., 1997). The effort and cost are justified it if it is possible to generate a genome map for a tree species that results to be an appropriate candidate for domestication and cultivation. Though many molecular markers are constantly being generated, there are few which are co-dominant in nature. Hence, RFLP markers are still considered as the molecular marker of choice, if quantitative trait locus mapping is the ultimate goal. The cost for RFLP mapping is difficult to estimate. Depending on the number of markers needed, this could vary between $5,000 (for 10 markers) to $50,000 for exhaustive search of RFLP markers on a genome.

3.2.7. Techniques for clone identification

VNTR have been the technique of choice for fingerprinting individual clones or ramets. M13 probes and Jeffreys' probes 33.6 and 33.15 have been used to identify cultivars (Sharon, *et al.*, 1995). In the past, isozymes have been used for germplasm identification. However, often they are not sensitive enough and VNTR have served well for this purpose. Other techniques such as SSRs (Simple Sequence Repeats) (used often for mammals particularly when scoring for 6 or more loci) are also being explored for the purpose of identification. The use of AFLP will have strong potential in this area of research. For the purpose of quick and low cost screening of individuals, or seed-lots, RAPD markers will also be useful. If a particular band is consistently variable between individuals (therefore expressing high levels of polymorphism), one could clone these bands, using RAPD or VNTR methods, sequence and design primers that would flank these variable regions on the genome. They can be used for checking seed lots prior to planting and to verify the correct maternal and paternal parents.

3.3. General consideration on the use of different techniques

MICROSATELLITES. Microsatellites are known for their high mutation rates; scores are made based on size differences between 2 alleles for a given individual. Therefore, microsatellite data can be used in many types of population studies including deduction of parentage, outcrossing rates and even pollen flow. A good example in a tropical tree would be the work of Dawson *et al*. (1997) on *Gliricidia sepium*. Using one microsatellite primer set, they found that with the 6 alleles (2 common and 4 “rare”) they could calculate the minimum distance for pollen flow in this insect pollinated tree. Similarly, the use of microsatellites was critical in the analysis of seed dispersal in the temperate tree, bur oak (*Quercus macrocarpa*). Dow and Ashley (1996) found that long distance seed dispersal could be more common than it has been previously reported.

Conversely, the high rate of mutation can be a hindrance when using microsatellite markers and researchers should be aware of other issues such as 1) scoring of gels, 2)

observation (or lack there of) of null alleles, 3) homoplasy and 4) the use of heterologous primers during cross species amplification.
A "known" allelic ladder should be used in all gels so that cross comparison of gels can be made accurately when scoring. Another problem is the production of "stutter" bands, which can cover or reduce the number of heterozygotes if one is not confident in scoring microsatellites. Typically, a primary allele band will be darker than a stutter band.
The use of cross-species microsatellite primers should be done with caution. The probability of null alleles (non amplification of one or more alleles) needs to be carefully analyzed. Often with non-focal species there will be base changes in the flanking region of the microsatellites particularly if the non-focal species is distantly related to the focal species (Callen *et al*., 1993).
The usefulness of microsatellite markers for many applications when studying population biology can be founded across many disciplines. From parentage and relatedness to migration and assignment tests and stock analysis for management purposes. These topics can all be addressed using microsatellite markers (Sunnucks, 2000). Many data analysis programs can be found on the internet and some of them are user friendly, such as GENEPOP (Raymond *et al*., 1995). More demanding is sometimes data collection. There are several mathematical models specific to microsatellite data. Two of the most common models are the IAM (Infinite Allele model) and SMM (step-wise mutation model). A good review of these and other models can be found in Goldstein and Pollack (1997) paper. Unlike, isozyme, RAPD and AFLP, allele genealogy can be found in the sequences of the microsatellites. A quick comparison of the sequences obtained from a known source and the sequences of the microsatellite from an organism of choice can be made to deduce the history of the alleles. However, the complexity of data resulting from the use of microsatellite on non-focal species would reduce the possibility of comparison of data across many studies.

AFLP. Data analysis issues for AFLP are very similar to those of RAPD (see above sections). An extra data scoring issue is the co-segregation of two or more sets of bands which can be either tightly linked or the second is generated by an internal restriction site. A check for Mendelian inheritance or evidence of incomplete linkage needs to be done on any AFLP data set. Like RAPD data, AFLP markers should be used on pedigrees whenever possible. Unlike RAPD data, AFLP markers are very robust and repeatable. For AFLP, the allele genealogy is rarely possible and comparison across studies, like RAPD, is minimal unless the same markers are used on the same individuals, population and/or species.

3.4. Summary on the applications of the techniques described

Population geneticists as well as forest geneticists are quickly embracing the new molecular techniques, such AFLP and SSR. These techniques are extensively used for the study of genetic diversity, QTL mapping, seed lot and ramet identification as well as population structure analysis. Since some of these techniques can be more labour intensive and expensive, costs and time involved need to be evaluated. Co-dominant markers such as SSR could be well considered particularly to study trees that are valuable candidates for domestication and cultivation.

4. Model forest organisms

Genetic markers are used not only to study genetic variations, using techniques such as microsatellites and AFLP. They are also applied for QTL (Quantitative Loci Traits) mapping and MAS (Marker Assisted Selection) applied to quantitative traits such as growth rate, quality of the wood and other environmental adaptive traits.

For trees such as *Ecalyputus* (Gaiotto *et al*., 1997) and Lobolly pine (D. Neale, pers. com.) genetics markers such as RAPD, AFLP and microsatellites have generated detailed genomic maps. For many crop plants such as rice, barley and tomato, genome maps are generated to study quantitative traits for disease resistance, drought tolerance and salt tolerance. In forestry, such progress is being made for traits such as frost, pest, and drought resistance with species such as lobolly pine, Sitka spruce and Douglas Fir. Using specific techniques such as cDNA sequencing, certain genes can be deduced. Partial cDNA sequences have been generated for loblolly pine (Whetten and Sederoff, 1992; Jermstad *et al*. 1998). Comparison of these cDNA sequences generated for loblolly pine showed high similarity with those genes found in angiosperm species (Jansson and Gustafsson 1991). Gene families from the ADH group (Alcohol dehydrogenase) are fairly common and two gene families have been identified. Isolation of mRNA during certain growth seasons or under environmentally challenged conditions can been carried out and genes can be identified using the method of cDNA sequencing. This method is laborious and costly and isolation of mRNA can be challenging under inexperienced hands. After the conversion of mRNA to cDNA, these pieces of DNA are then cloned and extensive probing with other known genes or “brute” sequencing ensues.

Conifers generally exhibit wind dispersal of pollen and seed, whereas many tropical trees usually have quite different pollination mechanisms such as insect or bird vectors. Of significant consequences is the understanding of micro-environmental selection for regeneration of forest trees (Epperson, 1992). Important genetic variation is often contained within a population and if these populations show micro-environmental adaptation, then choosing only those trees that have high trait value may not be practical. The subsequent replanting of "superior" trees may not be practical if the matching of micro-environmental conditions is required to induce these superior traits. The need of detailed knowledge of population microstructure would be required to assess the types of genetic variation necessary for the regeneration of a forest stand.

Aside from the obvious reasons of finding potential genes or assessing spatial structure, the study of biosystematics is essential for understanding the history of a species. Specific tree breeding programs particularly in search of a hybrid species that take the positive traits from two parents would be another reason for intensive research programs in forest genetics. Successful hybridization program does require knowledge of similarities between species involved and the more closely related the species, the more successful the program. The assay of bio-systematic characters would involve genotyping of sample individuals from a population. Potentially the most accurate character for analysis of relationships would be the use of nucleotide sequences. By focusing on DNA sequences that are not under natural selection such as introns between exons of genes on noncoding regions of the genome, these nucleotide changes would be ideal characters for studying the evolutionary changes leading to differentiation of species. Other markers such as RFLP could also be used as bio-systematic characters. In conjunction with data collection, the analysis of these characters requires understanding of the method chosen.

There are reasons that would deter the use of labor-intensive sequencing genetic markers. Firstly, the limitation of certain genetic markers such as RFLP and ribosomal DNA to identify above species level could limit the use of these markers for subspecies level or natural hybrids. With techniques such as EPICs (Exon Prime Intron Crossing) nucleotide sequences, this problem could possibly be resolved. Secondly, the cost of sequence techniques ($5.00 to $8.00 per sample) could be a deterrent. The other related limiting factors would be the need for technically trained personnel. Lastly, sequencing molecular analyses could be limiting due to the cost of the required software and the lack of user-friendly programs particularly for those that are new to the field of molecular genetics.

5. Cost-benefit analysis of alternative techniques

The table below is based on the knowledge of a given population size. For all the markers with the exception of sequences, there should be a minimum population of 30 individual per population. A pilot study should be done to deduce the amount of sampling needed to answer the question of genetic diversity for a given species. The symbol ($) is equivalent to the relative cost of data collection from $ = low to $$$$ = high.

Marker Choice **(Type of marker)**	**Single/Multi Loci** **(Allele number)**	**Small amount of tissues** (wet weight < 2 grams)	**Large amount of tissues** (wet weight > 2 grams)
Isozyme (Codominant)	Single (one to five)	√ ($)	√ ($)
RAPD **(Dominant)**	Multiple (up to two)	√ ($)	√ ($)
Microsatellite **(Codominant)**	**Single** **(two to many)**	**√** **($$)*** *** after initial cost of development ($$$$)**	**√** **($$)*** *** after initial cost of development ($$$$)**
AFLP **(Dominant)**	**Multiple** **(up to two)**		**√** **($$$)**
RFLP **(Codominant)**	**Single** **(up to four)**		**√** **($$$)**

Minisatellite **(Dominant)**	**Multiple** **(many)**		**√** **($$$)**
VNTR **(Dominant)**	**Multiple** **(many)**		**√** **($$$)**
Sequences **(Codominant)**	**Single** **(up to four)**	**√** **($$$$)**	**√** **($$$$)**

6. Conclusions

The choice of genetic markers should always be made based on the biological question at hand. For "wild" species with unknown genetic information, there will always be a need for basic data such as the level of heterozygosity, percent of polymorphic markers within and between populations. Initial sampling should be done on the range of populations to answer questions of genetic diversity across the species range. For trees of economic value, questions of mating and breeding system could be addressed with various genetic markers now made available. Like genetic model organisms (such as *Drosophila*, *Mimulus* and Mouse), there are tree models that we can use to ask deeper questions and to investigate properties of gene trait selection and genetic modification to produce desired results such as drought, pest and disease resistance. The need for genetic markers will always remain for any type of population genetics study and the growing field of forest genetics is no exception.

References

Aitken, S. N. and Libby, W. J. 1994 Evolution of the pygmy-forest edaphic subspecies of *Pinus contorta* across an ecological staircase Evol. 48(4):1009-1019.

Ali, I. F., Neale, D. B. and K. A. Marshall 1991 Chloroplast DNA restriction fragment length polymorphism in Sequoia sempervirens D. Don Endl., *Pseudotsuga menziesii* (Mirb.) Franco, *Calocedrus decurens* (Torr.), and *Pinus taeda* L. Theor. Appl. Genet. 81:83-89.

Amos, W. and Rubinsztein, D. C. 1996 Microsatellites are subject to directional evolution. Nat. Genet. 12:13-14.

Ayad, W. G., Hodgkin, T., Haradat, A. and V. R. Rao 1995 Molecular genetic techniques for plant genetic resources. Report of an IPGRI workshop, 9-11 October 1995, Rome, Italy. International Plant Genetic Resources Institute, Rome, Italy.

Becker, J. and M. Heun. 1995 Barley microsatellites: allele variation and mapping. Pl. Mol. Biol. 27:835-845.

Becker, J., Vos, P., Kuiper, M., Salamini, F. and M. Heun, 1995 Combined mapping of AFLP and RFLP markers in barley. Mol. Gen. Genet. 249:65-73.

Berg, E. E. and J. L. Hamrick 1997 Quantification of genetic diverity at allozyme loci. Can. J. For. Res. 27:415-424.

Bodenes, C., Joandet, S., Laigret, F. and A. Kremer 1997 Detection of genomic regions differentiating two closely related oak species *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. Heredity 78:433:444.

Botstein, D., White, R.L., Skolnick, M.H. and Davis, R.W. 1980. Construction of genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:641-656.

Callen D. F., Thompson, A.D., Shen Y., Phillips H.A., Richards, R.I. Mulley J.C., and Sutherland G.R. 1993 Incidence and origin of “null” alleles in the (AC)n microsatellite markers. Am J Hum Genet 52:922-927.

Chase, M. W. and H. H. Hills 1991 Silica gel: An ideal material for field preservation of leaf samples for DNA studies. Taxon 40:215-220.

Chase, M., Kesseli, R. and B. Kamaljit 1996a Microsatellite markers for population and conservation genetics of tropical trees. Amer. J. Bot. 83(1):51-57.

Chase, M. R., Moller, C., Kessell, R. and K. S. Bawa 1996b Distant gene flow in tropical tree. Nature 383:398-399.

Clark, A.G., Lanigan C.M.S. 1993 Prospects for estimating nucleotide divergence with RAPDs. Molecular Biology and Evolution 10:1096-1111.

Cnops, G. de Boer, B., Gerats, A. van Montagu, M. and M. van Lijsebettens. 1996. Chromosome landing at the Arabidopsis TORNADO1 locus using an AFLP based strategy. Mol. Gen. Genet. 253:32-41.

Condit, R. and S. P. Hubbell. 1991 Abundance and DNA sequence of two base repeat regions in tropical tree genomes. Genome 34: 66-71.

Cruzan, M. B. 1998 Genetic markers in plant evolutionary ecology. Ecol. 79(2):400-412.

Dawson, I.K., Waugh, R., Simons, A.J. and Powell, W. 1997 Simple sequence repeats provide a direct estimate of pollen-mediated gene dispersal in the tropical tree *Gliricidia sepium*. Mol. Ecol. 6:179-183.

De La Cruz, M., Whitkus, R. and M. Mota-Bravo 1995 Tropical tree DNA isolation and amplification. Mol. Ecol. 4:787-789.

Demesure, B., Sodzi, N. and R. J. Petit 1995 A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol. Ecol. 4:129-131.

Devey, M. E., Delfino-Mix, A., Kinloch, B. B. (Jr.) and D. B. Neale 1995 Random amplified polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. Proc. Natl. Acad. Sci. USA 92:2066-2070.

Dow, B.D. and Ashley, M.V. 1996 Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Mol. Ecol. 5:615-627.

Dumolin-Lapegue, S, Pemonge, M. H. and R. J. Petit 1997 An enlarged set of consensus primers for the study of organelle DNA in plants. Mol. Ecol. 6:393-397.

Echt, C. S. and P. May-Marquardt 1997 Survey of microsatellite DNA in pine. Genome 40:9-17.

Echt, C. S., May-Marquardt, P., Hseih, M., and R. Zahorchak 1996 Characterization of microsatellite markers in eastern white pine. Genome 39;1102-1108.

Edwards, A., H.A. Hammond, L. Jin, C.T. Caskey and R. Chakraborty. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241-253.

El-Kassaby, Y. A. and A. E. Szmidt 1988 The use of restriction analysis of chloroplast DNA in classifying hybrid spruce seedlots. IN: Proc. Frans. Kempa Symp., Molecular genetics of forest trees. Umea, Sweden. Report No. 8:67-88.

El-Kassaby, Y. A. and Ritland, K. 1996a Genetic variation in low elevation Douglas-fir of British Columbia and its relevance to gene conservation. Biodiver. Conserv. 5:779-794.

El-Kassaby, Y. A. and Ritland, K. 1996b Impact of selection and breeding on the genetic diversity in Douglas-fir. Biodiv. Conserv. 6:795-813.

Epperson, B.K. 1992 Spatial structure of genetic variation within populations of forest trees. New Forests 6:257-278.

Ferris, C., Oliver, R. P., Davy, A. J. and G. M. Hewitt 1995 Using chloroplast DNA to trace postglacial migration routes of oaks into Britain. Mol. Ecol. 4:731-738.

Gaiotto, R. A., Bramucci, M. and D. Grattapaglia 1997 Estimatin of outcrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. Theor. Appl. Genet. 95:842-849.

Goldstein, D.B. and D.D. Pollack 1997 Launching microsatellites: A review of mutation processes and methods of Phylogenetic inference. J. Hered. 88:335-342.

Hamrick J. L. and D. A. Murawski 1991 Levels of allozyme diversity in populations of uncommon neotropical tree species. J Trop. Ecol. 7:395-399.

Hamrick, J. L. and Godt, M. J. W. 1989 Allozyme diversity in plant species. IN: Differentiation patterns in higher plants (K. Urbanska, ed.) pp. 53-67. New York: Academic Press.

Heun, M., Schafer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, G. and F. Salamini, 1997 Site of einkorn wheat domestication identified by DNA fingerprinting. Science 278:1312-1314.

Jansson, S. and P. Gustafsson 1991 Evolutionary conservation of the chlorophyll a/b-binding proteins: cDNAs encoding Type I, II and III LHC I polypeptides from the gymnsperm Scots pine. Mol. Gen. Genet. 229:67-76.

Jarne, P. and P. J. L. Lagoda 1996 Microsatellites, from molecules to populations and back. TREE 11(10):424-429.

Jeffreys, A.J., Wilson, V. and S. L. Thein. 1985 Individual-specific 'fingerprinters' of human DNA. Nature 316:76-79.

Jermstad, K.D., Bassoni, D.L., Kinlaw, C.S. and D.B. Neale 1998 Partial DNA

sequencing of Douglas-fir cDNAs used for RFLP mapping. T.A.G. 97(5/6):771-776.

Jobes, D.V., Hurley, D.L. and L. B. Thien 1995 Plant DNA isolation: a method to efficiently remove
polyphenolics, polysaccharides, and RNA. Taxon 44:379-386.

Kamalay, J. C. and D. W. Carey 1996 Using DNA markers to identify American Elm. Amer. Nursery. April 15: 56-57.

Karp, A., Kresovich, S., Bhat, K. V., Ayad, W. G. and T. Hodgkin 1997 Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No.2. International Plant Genetic Resources Institute, Rome, Italy.

Keim, P., Schupp, J. M., Travis, S. E., Clayton, K., Zhu, T., Shi, L., Ferreira, A. and D. M. Webb, 1997 A high-density soybean genetic map based on AFLP markers. Crop Sci. 37:537-543.

Kostia, S., Varvio, S-l, Vakkair, P. and P. Pulkkinen 1995 Microsatellite sequences in a conifer, *Pinus sylvestris*. Genome 38:1244-1248.

Kremer, A., Petit, R.J. and Pons, O. 1998 Measures of polymorphism within and among populations. In: Molecular Tools for Screening Biodiversity. Eds. Karp, A., Isaac, P.G, and Ingram, D.S. (Chapman & Hall, London). Pp. 301-311.

Lewontin, R.C. and J. Hubby. 1966 A molecular approach to the study of genic heterozygosity in natural populations. II Amounts of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54:595-609.

Li, P. and Adams, W. T, 1989 Range-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). Can. J. For. Res. 19:149-161.

Liu, Z. and Knowles, P. Patterns of allozyme variation in tamarack (*Larix laricina*) from northern Ontario. Can. J. Bot. 69:2468-2474.

Loveless, M.D. 1992 Isozyme variation in tropical trees:patterns of genetic organization. New Forests 6:67-94.

Lynch, M., Crease, T.J. 1990 The analysis of population survey data on DNA sequence variation. Mol. Biol. Evol. 7:377-394.

Ma, A.Q., Roder, M. and M.E. Sorrells. 1996 Frequencies and sequence characteristics of de-, tri-, and tetra-nucleotide microsatellites in wheat. Genome 39:123-130.

Mackill, D. J., Zhang., Z., Redona, E. D. and P. M. Colowit, 1996 Level of polymorphism and genetic mapping of AFLP markers in rice. Genome 39:969-977.

Marsan, P. A., Castiglioni, P., Fusari, F., Kuiper, M. and M. Motto 1998 Genetic diveristy and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor. Appl. Genet. 96:219-227.

Meksem, K., Leister, D., Peleman, J., Zabeau, M., Salamini, F. and C. Gebhardt, 1995 A high-resolution map of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers. Mol. Gen. Genet. 249:74-81.

Millar, C. I. , Strauss, S. H. and Conkle, M. T. and Westfall, R. D. 1988 Allozyme differentiation and biosystematics of the Californian closed-cone Pines (*Pinus* subsect. *Oocarpae*)

Mitton, J. B. 1983 Conifers. In Isozymes in plant genetics and breeding. Part B. Edited by S. d. Tanksley and T. J. Orton. Elservier Science Publishers, Amsterdam.

Moran, G. F. and Adams, W. T. Microgeographical patterns of allozyme differentiation in Douglas-fir from Southwest Oregon. Forest Sci. 35 (1):3-15.

Mullis, K. B., Faloona, F., Scharf, S. J., Saiki, R. K., Horn, G. T., and Erlich, H. A. 1986 Cold Spring Harbor Symp. Quant. Biol. 51:263-273.

Murphy, R.W., Sites, J.W.(Jr.), Buth, D.G. and Haulfer, (C.H.) 1996 Protein: Isozyme Electrophoresis. In Molecular Systematics (2nd edn.) Eds: Hillis, D.M., Mortiz, C., and Mable, B.K. Sinauer Assoc. (Massachusetts) pp.51-120

Nandi, S., Subudhi, P. K., Senadhira, D., Manigbas, N.L., Sen-Mandi, S. and N. Huang, 1997 Mapping QTLs for submergence tolerance in rice by AFLP analysis and selective genotyping. Mol. Gen. Genet. 255:1-8.

Neale, D. B. and Adams, W. T. 1985 The mating system in natural and shelterwood stands of Douglas-fir. Theor. Appl. Genet. 71:201-207.

Neale, D.B., Devey, M.E., Jermstad, K.D., Ahuja, M.R., Alosi, M.C. and Marshall, K.A. 1992. Use of DNA makers in forest tree improvement research. New Forests 6:391- 407.

Neale, D. B., Marshall, K. A. and D. E. Harry 1991 Inheritance of chloroplast and mitochondrial DNA in incense-cedar. Can. J. For. Res. 21:717-720.

Neale, D. B., Marshall, K. A. and R. R. Sederoff 1989 Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. Proc. Natl. Acad. Sci. 86:9347-9349.

Neale, D. B. and R. R. Sederoff 1989 Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. Theor. Appl. Genet. 77:212-216.

Neale, D. B. , Wheeler, N.C. and R. W. Allard 1986 Paternal inheritance of chloroplast DNA in Douglas fir. Can J. For. Res. 16:1152-1154.

Nesbitt, K. A., Potts, B. M. Vaillancourt, R. E., West, A. K. and J. B. Reid 1995 Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). Heredity 74:628-637.

Pakniyat, H., Powell, W., Baird, E., Handley, L. L., Robinson, D., Scrimgeour, C. M., Nevo, E., Hackett, C. A., Caligari, P. D. S. and B. P. Forster, 1997 AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. Genome 40:332-341.

Perera, L., Russell, J. R. , Provan, J., McNicol, J. W. and W. Powell 1998 Evaluating genetic relationships between indigenous coconut (*Cocos nucifera* L.) accessions from Sri Lanka by means of AFLP profiling. Theor. Appl. Genet. 96:545-550.

Pfeiffer, A., Olivieri, A.M. and M. Morgante. 1997 Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.). Genome 40:411-419.

Pita, J. M., Perez-Garcia, F., Escudero, A. and C. dela Cuadra 1998 Viability of *Aven sativa* L. seeds after 10 years of storage in base collection. Field Crops Res. 55:183-187.

Plomion, C., Yani, A. and A. Marpeau 1996 Genetic determinism of δ3-carene in maritime pine using RAPD markers. Genome 39:1123-1127.

Powell, W., Morgante, M., Doyle, J. J., McNicol, J. W., Tingey, S. V. and A. J. Rafalski. 1996 Genepool variation in genus *Glycine* subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. Genetics 144: 793-803.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G. G. and J. A. Rafalski 1995 Polymorphic simple sequence repeat regions in chloroplast genomes: Applications to the population genetics of pines. Proc. Natl. Acad. Sci. USA 92:7759-7763.

Protopapadakis, E. E. 1987 Identification by isoenzymes of five cultivars of *Citrus medica* grafted on four rootstocks. J. Hort. Sci. 62(3): 413-419.

Qi, X. and P. Lindhout, 1997. Development of AFLP markers in barley. Mol. Gen. Genet. 254:330-336.

Rajora, O. P. 1989 Marker allozyme genes and alleles for differentiation *of Populus deltoides, P. nigra, P. maximowiczii*, and their interspecific hybrids. Can. J. Bot. 68:990-998.

Ramser, J, Weising, K., Lopez-Prealta, C., Rerhalle, W., Terauchi, R. and G. Kahl 1997 Molecular marker based taxonomy and phylogeny of Guinea yam (*Dioscorea rotundata - D. cayenensis*). Genome 40:903-915.

Raymond, , M. and Rousset, F. 1995 GENEPOP (version 1.2): apopulation genetics software for exact tests and ecumenicism. J. Hered. 86:248-249.

Remington, D.L., Whetten, R.W., Liu, B.-H and D.M. O'Malley 1999 Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. T.A.G. 98:1279-1292.

Ritland, C. and K. Ritland 2000 DNA-Fragment markers in Plants. In: Molecular Methods in Ecology. Ed. A. J. Baker (Blackwell Sci., Oxford) pp.208-234.

Rogers, S. O. and A. J. Bendich 1985 Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Pl. Mol. Biol. 5:69-76.

Rogstad, S. H. 1993 Surveying plant genomes for variable number of tandem repeat loci. Method Enzymol. 224:278-294.

Rouppe van der Voort, J. N. A.M., P., van Eck, H. J., Folkertsma, R. T., Hutten, R. C. B., Draaistra, J., Gommers, F. J., Jacobsen, E., Helder, J. & J. Baker 1997 Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol. Gen. Genet. 255:438-447.

Russell J. R., Fuller, J. D., Macaulay, M., Hatz, B.G. Jahoor, A. , Powel, W. and R. Waugh 1997 Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95:714-722.

Saiki, R. K., Scharf, S., Faloon, F., Mullis, K., Horn, G., Erlich, H. A., and Arnheim, N. 1985 Enzymatic amplification of beta-globin sequences and restriction site analysis diagnosis of sickle cell anemia. Science 230:1350.

Schmidt, T. and J. S. Heslop-Harrison 1996 The physical and genomic organization of microsatellites in sugar beet. Proc. Natl. Acad. Sci. USA 93:8761-8765.

Sharon, D., Adato, A., Mhameed, S. and U. Lavi 1995 DNA fingerprints in plants using simple-sequence repeat and minisatellite probes.

Smith, D. N. and M. E. Devey 1994 Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37:977-983.

Sokal, R.R. ans Rohl,F.J. (1981) Biometry 2n edn. W.H. Freeman, San Francisco.

Southern, E. M. 1976 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Sunnucks P. 2000 Efficient genetic markers for population biology. TREE 15:199-203.

Sutherlands, G. R. and Richards, R. I. (1995) Simple tandem DNA repeats and human genetic disease. Proc. Natl. Acad. Sci. 92:3636-3641.

Sutton, B. C. S. , Flanagan, D. J. and Y. A. El-Kassaby 1991a A simple and rapid method for estimating representation of species in spruce seedlots using chloroplasts DNA restriction fragment length polymorphism. Silvae Gene. 40:119-123.

Sutton, B. C. S., Flanagan, D. J., Gawley, J. R., Newton, C. H., Lester, D. T. and Y. A. El-Kassaby 1991b Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. Theor. Appl. Genet. 82:242-248.

Swofford , D.L. and Selander, R.B. 1981 BIOSYS-1:a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. Journal of Heredity 72:281-283.

Swofford, D.L., Olsen, G.J., Waddell, P.J. and Hillis, D.M. 1996 Phylogenetic Inference. In Molecular Systematics (2^{nd} edn.) Eds: Hillis, D.M., Mortiz, C., and Mable, B.K. Sinauer Assoc. (Massachusetts) pp.407-514.

Szmidt, A. E. , El-Kassaby, Y. A., Sigurgeirsson, A., Alden, T., Lindgren, D. and J-E. Hallgren 1988 Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression using restriction analysis of chloroplast DNA. Theor. Appl. Genet. 76:841-845.

Szmidt, A. E., Alden, T. and J-E. Hallgren 1987 Paternal inheritance of chloroplast DNA in *Larix*. Pl. Mol. Biol. 9:59-64.

Taberlet, P., Gielly, L., Pautou, G. and J. Bouvet 1991 Universal primers for amplification of three non-coding regions of chloroplast DNA. Pl. Mol. Biol. 17:1105-1109.

Tanksley, S. D. and T. J. Orton (eds.) 1983 Developments in plant genetis and breeding 1: Isozymes in plant genetics and breeding. Part A. Elsevier Science Publishers B.V., New York, New York.

Terauchi, R. 1994 A polymorphic microsatellite marker from the tropical tree Dryobalanops lanceolata (Dipterocarpaceae). Jpn. J. Genet. 69:567-576.

Terauchi, R., and A. Konuma. 1994 Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. Genome 37:794-801.

Tohme, J., Gonzalez, D.O., Beebe, S. and M. C. Duque, 1996 AFLP analysis of gene polls of a wild bean core collection. Crop Sci. 36:1375-1384.

Travis, S. E., Maschinski, J. and P. Keim, 1996 An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically endangered plant, using AFLP markers. Mol. Ecol. 5:735-745.

Tulsieram, L. K., Glaubitz, J. C., Kiss, G. and J. E. Carlson 1992 Single tree genetic linkage mapping in conifers using haploid DNA from megagametophytes. Biotech. 10(6):686-690.

Van de Ven, W. T. G. and R. J. McNicol 1995 The use of RAPD markers for the identification of Sitka spruce (*Picea sitchensis*) clones. Heredity 75:126-132.

VanToai, T. T., Peng, J. and S. K. St. Martin, 1997 Using AFLP markers to determine the genomic contribution of parents to populations. Crop Sci. 37:1370-1373.

Vendramin, G.G. and B. Ziegenhagen 1997 Characterisation and inheritance of polymorphic plastid microsatellites in *Abies.* Genome 40:857-864.

Vendramin, G.G., Lelli, L. Rossi, P. and M. Morgante 1996 A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol. Ecol. 5:595-598.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van der Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman J., Kuiper, M., and M. Zabeau, 1995 AFLP: a new technique for DNA fingerprinting. Nuc. Acid Res. 23:4407-4414.

Waugh, R., Bonar, N., Baird, E., Thomas, B., Graner, A., Hayes, P. and W. Powell, 1997 Homlogy of AFLP products in three mapping populations of barley. Mol. Gen. Genet. 255:311-321.

Whetten, R. W. and Sederoff, R. R. 1992 Phenylalanine ammonia-lyase from Loblolly Pine. Phant Physiol. 98:380-386.

Williams, J. G. K., Kubelik, A. R., livak, K. J., Rafalski, J. A., and S. V. Tingey 1990 DNA polymorphoisms amplified by arbitrary primers are useful as genetic markers. Nuc. Acids Res. 18(22):6531-6535.

Wu, K-S. and S. D. Tanksley. 1993 Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol. Gen. Genetics 241:225-235.

Yeh, F. C. H. and O'Malley, D. 1980 Enzyme variations in matural populations of Douglas-fir, *Pseudotsuga menziesii* (Mirb.) Franco, from British Columbia. 1. Genetic variation patterns in coastal populations. Silvae Genetica 29:83-92.

Zhu, J. Gale, M. D., Quarrie, S., Jackson, M. T. and G. J. Bryan, 1998 AFLP markers for the study of rice biodiversity. Theor. Appl. Genet. 96:602-611.